DOCUMENT RESUME

ED 163 957 IR 006 583

AUTHOR Braden, Roberts A.; Walker, Alice D.
TITLE Evaluation Is. . .
PUB DATE 78
NOTE 14p.; Paper presented at the Annual Meeting of the Association for Educational Communications and Technology (Kansas City, Missouri, April 1978)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Educational Accountability; Educational Alternatives; *Educational Assessment; *Evaluation Methods; *Evaluative Thinking; Measurement Techniques; Relevance (Education); *State of the Art Reviews

ABSTRACT

The synthesis of historical definitions presented in this paper was designed to trace the development of the process of evaluation, showing how modern educational evaluation has evolved and identifying some points of confusion. The importance of the integration of evaluation into the overall educational process is emphasized. (Author/RAO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED163957

EVALUATION IS . . .

by

Roberts A. Braden
Head, Instructional Development Division
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

and

Alice D. Walker
Editorial Assistant
Instructional Development Division
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Alice D. Walker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR006 583

Evaluation is. . . . Obviously, evaluation is many things to many people, but it is a word we all use, and therefore we all think we know what it is. While preparing for a presentation at the national AECT convention in Kansas City, the authors found that evaluation is a frequently misunderstood term. An intensive review of current literature reinforced the belief that there is a wide divergence of opinion surrounding what might be considered basic terminology. Evaluation has so many different meanings to different people that when we say it, hear it, write it, or read it, we cannot be sure that the other person is using our meaning. In this context, and so that we may all start on common ground, consider that The American Heritage Dictionary (1971) distinguishes among six synonyms (estimate, appraise, assess, assay, evaluate, and rate) by stating, "Evaluate implies considered judgment in setting a value on a person or a thing." The key words in that definition are judgment and value.

It is the area of judgment that seems to produce the biggest hangups. No one wants to be the victim of judgment, final or temporary. And, few people enjoy the onerous task of passing judgment if it is upon another person. Thus, a faculty member may assign grades but vehemently object to student evaluation. At the same time, a student may complain about the course and grade received, but be reluctant to complete an evaluation form. It seems clear to the authors that herein lies mischief. The tendency for the acts and issues of evaluation to become emotionally charged only complicates the problem of the diverse meanings given to the term.

Much of the current furor being raised in academic circles concerning educational evaluation is due in large measure to a communication gap. Although a multitude of texts and articles have been produced, Kunz (1978) declares as recently as February of this year that "many professors' objections to evaluation simply arise from confusion over what it is." Much of the confusion, we believe, is derived from a reluctance on the part of academe to make the transition from knowledge to action. New theories and more advanced models, important as they are to the development of the field, will not accomplish the task of diminishing the growing antagonism toward this important segment of the educational process. Rather than confound the problem further, we offer a simple synthesis of historical definitions. Taken out of context, the quotations we cite can hardly be expected to reflect the total philosophy of their authors. It is our purpose, however, to use these definitions to trace the development of the process of evaluation, to show how modern educational evaluation has evolved, and to attempt to identify some of the points of confusion. Perhaps in this decontaminating, demystifying process we shall succeed in offering a "new" definition that will meet the needs of today's instructional setting.

Evaluation is, has been, and evermore shall be. Evaluation began in the beginning. Thumbing quickly through the millenia, we find reports that Chinese emperors used proficiency tests as early as 2200 B.C. The Bible in both Testaments stresses the importance of judgment. Our literature, from the Greek classics through Chaucer to the present, has

focused upon the testing of men and things. The production of a "Scale-Book" by the Reverend George Fisher in 1864 has been called "the starting point of measurement in evaluation" (Merwin, 1969).

What then is the starting point for educational evaluation? Does it lie in the queries of Socrates or the student rejections of incompetent faculty at the primitive University of Bologna? Maybe, but for our present concerns we would do better to begin with Ralph Tyler, recognized by many as the father of modern educational evaluation. In his classic text, Tyler (1949) defines evaluation as "the process of determining to what extent the educational objectives are actually being realized by the program of curriculum and instruction."

Tyler's emphasis on objectives led to the publishing of many "how-to" books. As educators scurried to write well-defined objectives, psychometrists sought new methods of measurement. At the same time, national interest in science created a need for more theoretical application of statistics. Behavioral psychologists, experimental researchers, and cost accountants demanded "scientific" data to be plugged into carefully structured models. Whereas in 1961 evaluation was termed "closely related to measurement" (Thorndike & Hagen, 1961), by the early seventies evaluation became "roughly synonymous with measurement" (Worthen, 1973).

As applications of educational evaluation became increasingly complex, efforts were made to simplify the process through the design of general models. Stufflebeam's (1971) basic context, input, process, product (CIPP) model was gradually expanded by other scholars, however,

until the number of boxes (more than 40 in Anderson & Flores, 1973), circles, hexagons, and arrows began to achieve a dizzying effect. Borich (1974) writes

> The field of educational evaluation has been increasingly complicated in past years by the growth and development of many evaluation models. Some of these models may only confuse evaluators by suggesting different approaches and by using diverse terminologies purported to be applicable to the same kinds of problems.

Borich carefully summarizes five specific models as an aid to uncomplicating the use of models by educators.

The continuing development of more highly sophisticated measurement techniques led some writers to use such disparaging phrases as "evaluation is a much overemphasized sacred cow" (Knowles, 1970). Educators were embarassed and disturbed because much of the information collected was useless--it covered limited areas and arrived too late for appropriate changes. Worthen's (1973) description of evaluation as "something which is usually inserted after the program has been implemented" reflects this general dissatisfaction.

More and more, educators were becoming interested not only in collecting data but in <u>using</u> that information to form judgments and select among alternatives (TenBrink, 1974). The role of educators as decision makers assumed new importance (Borich, 1974). So important was educational accountability that Geis (1974) referred to evaluation as "the key to instructional effectiveness." Perhaps most significant was the change from the negative ego-shattering threat of enforced

evaluation to a more positive cooperative effort to seek out effectiveness and worth (Diamond, 1975).

Gradually, the emphasis on precise measurement began changing to a more humanistic point of view. Kaufman (1975) presents a simplified description of evaluation as "the process for determining what works and what doesn't work, and revising that which doesn't." Anastasiow (1976) amplifies the issue of improvement by defining evaluation as ". . . an active search for information for the improvement of the project; an ongoing process to improve something rather than to prove it; a process of designing and refining."

Still, the controversies linger. It does matter whose ox is being gored. The whole matter of evaluation can take on different overtones depending upon whether you are the evaluator, the evaluatee, or a spectator. It would be foolish to pretend that the relationship of the evaluation system to the reward structure is an unimportant connection. The day is not yet here and may never come when most faculty members will cheerfully agree to the assessment of anything in the teaching-learning process other than student performance.

A wide diversity continues to exist, even among practicing evaluators, as to the appropriate definition, scope, purpose, and methods for educational evaluation. In the words of Knowles (1970), "evaluation is the source of more confusion, frustration, and guilt among educators than any other aspect of their work." Even more discouraging is the pronouncement by Rippey (1973): "At the moment, there seems to be no evidence that evaluation, although the law of the land, contributes anything to educational practice other than headaches

for the researcher, threats for the innovators, and depressing articles for journals devoted to evaluation." A glance at some of the adjectives applied to evaluation in these journal articles will illustrate the point: formal, informal; subjective, objective; macro, micro; congruency, contingency; qualitative, quantitative; comparative, non-comparative; heuristic, algorithmic; determinate, indeterminate; obtrusive, unobtrusive; acquirement, accomplishment; illuminative, judicial, transactional, and, of course, formative and summative.

Although evaluation has been termed a basic part of the instructional process, it is, according to Bloom (1969), "a two-edged sword." For every advantage evaluation specialists may list, a contradictory statement is supplied by its detractors. (Figure 1) Certainly evaluation procedures provide data, statistics, and numbers, but to some people this information may create more confusion than light. A tenured academic department head may feel legally protected by a systematized evaluation policy, whereas a junior staff member striving for promotion may feel threatened and frustrated. Some individuals may be inspired to greater achievement, knowing that they will be measured for their success, while others will react emotionally with guilt in fear that their failures will be documented. On the other hand, "the fact that the professor is not held responsible for the results of his efforts dulls the edge of his motivation, removes the incentive to follow through" (Fort Hays Kansas State University, 1978). Long range planning may reduce costs and make time more productive, but administrators will criticize the waste of time and money devoted to apparent non-essentials.

EVALUATION

PROVIDES DATA	CREATES CONFUSION
REPRESENTS PROTECTION	REPRESENTS THREAT
REDUCES FRUSTRATION	INCREASES FRUSTRATION
PREVENTS FAILURE	DETECTS FAILURE
INSPIRES SUCCESS	PROMOTES GUILT
STIMULATES GROWTH	HAMPERS CREATIVITY
SAVES MONEY	INVOLVES EXTENSIVE EXPENDITURES
MAKES TIME MORE PRODUCTIVE	TAKES TIME FROM PRODUCTIVE WORK
SHARPENS AND ILLUMINATES	BLURS AND OBSCURES

We should not be misled by the dynamics of the point-counterpoint aspect of the controversy. On the one hand, if we want to know whether instruction is working, we must evaluate. On the other hand, if we don't want the evaluation tail to wag the educational dog, we must compromise. Both the points and the counterpoints are valid to those who make them. Not everyone shares the evaluator's enthusiasm for evaluation. We should thus be prepared to accept less resources applied to evaluation than we believe are optimal. While we adamantly believe that evaluation must find a place in the system, it seems appropriate to suggest that it be a welcome place.

Our most recent definitions of evaluation reflect the influence of the instructional development process wherein the evaluator is no longer an external critic but has become an active member of the ID team. At the beginning of the seventies a respected authority on the subject of evaluation, Daniel Stufflebeam (1971), proposed that "Evaluation is the process of delineating, obtaining, and providing useful information for judging decision alternatives." In a recent interview (Brandt, 1978), Stufflebeam speaks of evaluation as "a communication process, a political process: it requires management skills." It is probably in this area of effective communication that we in the field of instructional technology meet our greatest challenge. Kunz (1978) writes, "evaluation is just one part of the puzzle of improving teaching and learning." Since the improvement of teaching and learning is the prime concern of the instructional developer, and a major concern of all other educators, it follows that we must learn to communicate the evaluation process in a manner which all can comprehend.

What is it that we do when we evaluate? We first consider all that we know about the subject. We may or may not collect additional information and make appropriate assumptions to fill the voids. We then apply some set of standards, or terms of reference, or scale or measure in conjunction with the collected knowledge. But the process does not end there--based on the information gathered we need to complete the evaluation process by making a value judgment.

Perhaps it is the seeming dichotomy of empirical and theoretical methods that has led to some of the confusion and frustration. Kunz (1978) resolves the issue with the simple statement, "Measurement is a science, evaluation is an art." It would seem to us that this is an oversimplification. Evaluation is both a science and an art; it is a process that uses judgment as the crucible for mixing art and science, theory and fact, arbitrary criteria and random samples.

Evaluation is an ever-evolving process whose focus has changed dramatically in recent years. Concepts involving who is to be evaluated, what is to be evaluated, and how evaluations are to be made will continue to evolve as educational practices continue to change. As technical capabilities become more complex it is important that we do not lose sight of the human factor. A recent report from the Southern Regional Educational Board (1977) included the statement, "Not all aspects of an evaluation system can consist of totally objective measurement; many aspects legitimately rely on subjective views and professional judgments. . . ." We agree; evaluation is a subjective act which we attempt to perform as objectively as possible. As professional educators we need to reassure ourselves that we are in fact capable of applying

reason to data and assessing the past as a guide to the future.

What is evaluation? Is it a process, an exercise, an application of something, a basis or a foundation for something, a bridge to somewhere? Is it a method that allows us to extrapolate, to find the essence, to identify opportunity, to correct errors, to avoid disaster, to improve, to change. . .? We can do little better than to simplify by going back to the original Middle English, Old French, and Latin, and drawing strength from the Anglo Saxons, declare in a strong, positive manner, "Evaluation is the judgment of worth." If we are to focus on the improvement of teaching and learning, should we not also concentrate on determining the worth of our efforts?

BIBLIOGRAPHY

Anastasiow, N. J. and Hanes, M. L., "A Case Study of Evaluating a Follow Through Program," Viewpoints, Vol. 52, No. 4, July, 1976, pp. 36-61.

Anderson, D. A. and Flores, T. R., "Implementing Systematic Evaluation Within An Ongoing Educational Program," Educational Technology, Vol. XIII, No. 5, June, 1973.

Bloom, B. S., "Some Theoretical Issues Relating to Educational Evaluation," in EDUCATIONAL EVALUATION: NEW ROLES, NEW MEANS, The 68th Yearbook of the National Society for the Study of Education, Chicago: University of Chicago Press, 1969, pp. 26-50.

Borich, G. D. (ed.) EVALUATING EDUCATIONAL PROGRAMS AND PRODUCTS. Englewood Cliffs, New Jersey: Educational Technology Publications, 1974.

Brandt, R., "On Evaluation: An Interview with Daniel L. Stufflebeam," Educational Leadership, Vol. 35, No. 4, January, 1978, pp. 248-254.

Diamond, R. M., et al., INSTRUCTIONAL DEVELOPMENT FOR INDIVIDUALIZED LEARNING IN HIGHER EDUCATION. Englewood Cliffs, New Jersey: Educational Technology Publications, 1975.

Fort Hays Kansas State University, IMPROVING INSTRUCTION, "Great Teaching," 1978.

Geis, G. L., "Evaluation: Definitions, Problems and Strategies," in Knapper, C. K. et al., IF TEACHING IS IMPORTANT. Toronto: Clarke, Irwin, & Company, Ltd., 1974, pp. 8-41.

Kaufman, R., NEEDS ASSESSMENT: WHAT IT IS--AND HOW TO DO IT. San Diego: United States International University, 1975.

Knowles, M. S., THE MODERN PRACTICE OF ADULT EDUCATION. New York: Association Press, 1970.

Kunz, D., "Learning to Live With Evaluation," Change, Vol. 10, No. 2, February, 1978, pp. 10-11.

Merwin, J. C., "Historical Review of Changing Concepts of Evaluation," in EDUCATIONAL EVALUATION: NEW ROLES, NEW MEANS, The 68th Yearbook of the National Society for the Study of Education, Chicago: University of Chicago Press, 1969, pp. 6-25.

Morris, William (ed.), THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE. Boston: A. H. Publishing Co. & Houghton Mifflin Co., 1971.

Rippey, R. M. (ed.), STUDIES IN TRANSACTIONAL EVALUATION. Berkeley: McCutchan Publishing Corp., 1973.

Southern Regional Education Board, FACULTY EVALUATION FOR IMPROVED LEARNING. Atlanta, 1977.

Stufflebeam, D. I., et al., EDUCATIONAL EVALUATION & DECISION MAKING. Itasca, Illinois: F. E. Peacock Publishers, 1971.

TenBrink, T. D., EVALUATION: A PRACTICAL GUIDE FOR TEACHERS. New York: McGraw Hill Book Company, 1974.

Thorndike, R. L. and Hagen, Elizabeth. MEASUREMENT AND EVALUATION IN PSYCHOLOGY AND EDUCATION. New York: John Wiley & Sons, Inc., 1961.

Tyler, R. W., BASIC PRINCIPLES OF CURRICULUM AND INSTRUCTION. Chicago: The University of Chicago Press, 1949.

Worthen, B. R. and Sanders, J. R., EDUCATIONAL EVALUATION: THEORY AND PRACTICE. Worthington, Ohio: Charles A. Jones, 1973.